

Agenda Item 2.4: Strategy for Countywide Open Data

Strategic Advisory Council
April 2, 2010

Background: King County is planning to update our code directing agencies to *strive* to publish “High Value Data Set” datasets on-line, in an *open format*. This legislation will be modeled after federal [Data.Gov](http://www.data.gov) (<http://www.data.gov>) and similar efforts by other local jurisdictions with the goal of improving governmental transparency and enabling innovation in our region. Our approach will be to have a single county Web site for all data sets. Data sets must be appropriate for unrestricted public disclosure.

In **Exhibit 1** we have listed some of the issues surrounding open data that need to be considered, such as data protection and costs and benefits associated with this effort. Your advice on these issues would help us in developing a countywide approach.

In considering approach and information that could be made available thru this initiative, we have considered a phased approach and some potential data sets for early phases. They are provided in **Exhibit 2: List of Potential Data Sets to be included on the open data site**.

A text of proposed ordinance relating to establishing a requirement for all county agencies to publish high value data sets on a county Web site is included in **Exhibit 3: King County Proposed Ordinance 2010-0100**.

YOUR ACTION IS REQUESTED

1. **Provide questions/input in advance of the meeting.** If you have questions or input related to this subject, we would appreciate receiving your feedback in advance of the meeting. For questions please contact David Martinez; otherwise, please e-mail your input associated with this agenda topic to Maureen Weisser. Maureen will be collecting all input provided by the members. We will introduce the received feedback at the SAC meeting for discussion.
2. **SAC Action at the meeting:** The members will be asked to vote on:
 - a. Endorsing countywide approach
 - b. Publishing of government /citizen data sets to be organized and facilitated by the Office of County Chief Information Officer
 - c. Including strategic objective for Open Data in *Strategic Technology Plan 2009-2012*

Exhibit 1: Issues Surrounding Open Data

King County is planning to update its governing legislation to require internal data sets be shared with the community. This effort is in alignment with recent federal direction to make government more open and transparent. A side benefit of this transparency is that innovators in the community are able to utilize the information and make it more valuable to citizens. Examples of this are sending bus arrival times to your phone while you are waiting at a bus stop, or letting you see crime distribution over time for a neighborhood.

Given these goals, King County has identified several issues that could benefit from your perspective and advice. As we are analyzing how to best move towards open data, the following concerns and issues have been identified:

- How to ensure that Sensitive Data is protected and not made public
- Should we consider data as Intellectual Property or should it all be treated like public records?
 - If public records, should others be allowed to profit from use?
 - Should costs to provide open data be recovered through usage? Or is it a government service that our citizens are already entitled to?
- What are the true costs of open data in both staff time and resources?
 - What level of service should be provided for
 - Questions on what data means or how it is organized
 - Requests for additional data or customized data
 - Corrections to data
 - Data refresh requests
 - What level of infrastructure is needed
 - What data refresh rates make sense
- Might some data be misused (i.e. crime stats helping criminals to be more effective?)
- Should we focus on both benefits? Or focus on one first and then the other?
 - Transparency and accountability
 - Enabling community innovators
- Should we take a regional approach
 - Hosting others' data to improve innovation?
 - Hosting industry as well as governmental data?

Exhibit 2: List of Potential Data Sets to be included on the open data site

Potential Data Sets by Phase
March 17, 2010

The following are the likely “Phase I” and “Phase II” data sets to be published by the CIO’s office under the “recommended approach for open data”.

This list identifies the description of the data set, the source of the data, and an example of the type of data within the proposed data set. There are 17 data sets in Phase I, and an additional 15-25 more data sets in Phase II.

1. Phase I – Publishing existing simple datasets

- KC Employee dataset
 - i. Pulled from online directory database
 - ii. Employee name, position, contact info
- Permits dataset
 - i. Pulled from Permit Search database
 - ii. Parcel number, permit type, date, status, applicant
- Restaurant Inspection dataset
 - i. Pulled from restaurant inspection database
 - ii. Restaurant, inspection type, date, score, violation
- Bus Schedule dataset
 - i. Pulled from online bus schedule database
 - ii. Route, stops, schedule details
- Property Change History dataset
 - i. Pulled from assessor data
 - ii. Property details, history, dates
- Apartment Complex dataset
 - i. Pulled from assessor data
 - ii. Property location, descriptors
- Commercial Building dataset
 - i. Pulled from assessor data
 - ii. Property location, descriptors
- Condo Complex dataset
 - i. Pulled from assessor data
 - ii. Property location, descriptors

- Residential Buildings dataset
 - i. Pulled from assessor data
 - ii. Property location, descriptors
- Vacant Lot dataset
 - i. Pulled from assessor data
 - ii. Property location, descriptors
- Environmental Restrictions dataset
 - i. Pulled from assessor data
 - ii. Property details, restrictions, dates
- Legal Property Restrictions dataset
 - i. Pulled from assessor data
 - ii. Property details, restrictions, dates
- Parcel Definition dataset
 - i. Pulled from assessor data
 - ii. Property location, descriptors
- Real Property Sales dataset
 - i. Pulled from assessor data
 - ii. Property details, transaction history, dates
- Full Property History dataset
 - i. Pulled from assessor data
 - ii. Property details, history, dates
- Tax Account Value dataset
 - i. Pulled from assessor data
 - ii. Property details, tax history, dates
- Value History dataset
 - i. Pulled from assessor data
 - ii. Property details, valuation history, dates

2. Phase II – Publishing of more innovative datasets (the following examples to be further analyzed)

- Bus Location tracking
 - i. Pulled from existing data feeds
 - ii. Employee name, position, contact info
- KC Job Openings dataset
 - i. Pulled from online job posting

- ii. Job title, dept, salary, closing
- Jail Register data
 - i. Pulled from JILS search data
 - ii. Suspect, charges, dates
- County Development Projects data
 - i. Pulled from BRED data
 - ii. Projects, value, subcontractors, status
- County Vendor data
 - i. Pulled from finance data
 - ii. Vendors, projects, invoices, pay status
- Department/Division/Organization Budget data
 - i. Pulled from finance data
 - ii. County organizations, budget, current spending
- Pets data
 - i. Pulled from pets databases
 - ii. Pet types, owners, adoption dates
- Candidate Registration data
 - i. Data on candidates filing for elected office
 - ii. Candidates, position, date, filing details
- Crime Stats data
 - i. Data on crime stats, IF available
 - ii. Crimes, locations, dates

Exhibit 3: King County Proposed Ordinance 2010-0100

AN ORDINANCE relating to establishing a requirement for county agencies to publish high value data sets on a county web site; and adding a new section to K.C.C. chapter 2.16.

STATEMENT OF FACTS:

1. Many cities, including San Francisco, New York and Washington, D.C., have released public data to web and mobile application developers in order to allow for the development of applications that will benefit the public.
2. The federal government, as part of its Open Government Directive requires all federal agencies to release high-value data sets.
3. Data that resides in technology systems is different from information such as documents, emails and calendars, in that it is structured and can be used by other computer applications for analysis or new uses such as mapping.
4. Providing access to county data can lead to innovation in how the public interacts with government, resulting in social and economic benefits for the county.
5. King County has a long history of innovation and is home to hundreds of technology companies that are changing the way the world operates.

6. By bringing county data and King County's entrepreneurs together in one location, local industry can be stimulated, jobs created and government information shared with the public.

7. Citizens are the greatest asset of any government. There are many benefits to engaging our highly skilled workforce in the development of applications and increasing citizen access to King County information.

BE IT ORDAINED BY THE COUNCIL OF KING COUNTY:

SECTION 1. The intent of this ordinance is to make King County government more accessible, transparent and accountable, and create economic opportunity by making high-value data sets available to the public.

NEW SECTION. SECTION 2. There is hereby added to K.C.C. chapter 2.16 a new section to read as follows:

A.1. Each agency shall strive to publish on-line, in an open format, datasets under their authority. To the extent practicable, each agency shall publish at least one high value data set by June 1, 2010.

2. For the purposes of this subsection A:

a. "high-value data set" means a data set that contains information that can be used to increase government accountability and responsiveness, improve public knowledge of the agency and its operations, further the core mission of the county, create economic opportunity or respond to need and demand as identified through public consultation; and

b. "open format" means a format that is platform-independent, machine-readable and made available to the public without restrictions that would impede the reuse of the information, and that can be read, downloaded, indexed and searched by commonly used web search applications.

B. All data sets shall be published on a single county web site maintained by the office of information resource management and designed to allow the public easy access to the data. The web site shall be developed by June 1, 2010 and incorporate a means for the public to:

1. Give feedback on and assessment of the quality of the data; and
2. Provide input about which data to prioritize for publication on the web site;

C. The chief information officer shall develop procedures and policies to ensure that all data sets available on the web site are appropriate for unrestricted public disclosure. The released data shall not include the names of individuals nor any personal identifying information. The released information must not contain any information that is exempt under the Public Records Act, chapter 42.56 RCW.

D. Each agency submitting data shall identify a person to whom the office of information and resource management may refer questions and concerns from users of the county's data sets.

E. The chief information officer shall develop terms of use for the data.